Chem. Senses 33: 639–653, 2008
doi:10.1093/chemse/bjn032
Advance Access publication July 5, 2008

Relationships Between Molecular Structure and Perceived Odor Quality of Ligands for a Human Olfactory Receptor

Guenhaël Sanz[1], Thierry Thomas-Danguin[2], El Hassan Hamdani[3], Claire Le Poupon[1], Loïc Briand[2], Jean-Claude Pernollet[1], Elisabeth Guichard[2] and Anne Tromelin[2]

[1]Institut National de la Recherche Agronomique, Unité Minte de Recherche 1197 Neurobiologie de l'Olfaction et de la Prise Alimentaire, F-78352 Jouy-en-Josas, France and [2]Institut National de la Recherche Agronomique, Unité Minte de Recherche 1129 Flaveur Vision et Comportement du Consommateur, F-21000 Dijon, France
[3]Present address: Biotechnology Centre of Oslo, University of Oslo, Norway

Correspondence to be sent to: Guenhaël Sanz, INRA, UMR 1197 Neurobiologie de l'Olfaction et de la Prise Alimentaire, Biochimie de l'Olfaction et de la Gustation, F-78352 Jouy-en-Josas. e-mail: guenhael.sanz@jouy.inra.fr

Abstract

Perception of thousands of odors by a few hundreds of olfactory receptors (ORs) results from a combinatorial coding, in which one OR recognizes multiple odorants and an odorant is recognized by a specific group of ORs. Moreover, odorants could act both as agonists or antagonists depending on the OR. This dual agonist–antagonist combinatorial coding is in good agreement with behavioral and psychophysical observations of mixture perception. We previously described the odorant repertoire of a human OR, OR1G1, identifying both agonists and antagonists. In this paper, we performed a 3D-quantitative structure–activity relationship (3D-QSAR) study of these ligands. We obtained a double-alignment model explaining previously reported experimental activities and permitting to predict novel agonists and antagonists for OR1G1. These model predictions were experimentally validated. Thereafter, we evaluated the statistical link between OR1G1 response to odorants, 3D-QSAR categorization of OR1G1 ligands, and their olfactory description. We demonstrated that OR1G1 recognizes a group of odorants that share both 3D structural and perceptual qualities. We hypothesized that OR1G1 contributes to the coding of waxy, fatty, and rose odors in humans.

Key words: 3D-QSAR, odor quality, odorant detection, odorant structure

Introduction

Humans are able to detect and discriminate myriads of structurally diverse odorants through their interaction with olfactory receptors (ORs) embedded in the plasma membrane of olfactory sensory neurons (OSN). ORs belong to the 7-transmembrane domain G protein–coupled receptor family and are encoded by an exceptionally large multigene family. Analysis of the human genome draft sequences has revealed approximately 380 potentially functional genes (Glusman et al. 2001; Zozulya et al. 2001; Malnic et al. 2004; Niimura and Nei 2005, 2007). It is well accepted that the perception of thousands of odors by a few hundreds of ORs results from a combinatorial coding, in which one OR recognizes multiple odorants and different odorants are recognized by different combinations of ORs (Duchamp-Viret et al. 1999; Malnic et al. 1999). Moreover, there is a growing body of evidence that odorants could act both as agonists or antagonists depending on the OR (Duchamp-Viret et al. 2003; Araneda et al. 2004; Oka et al. 2004; Sanz et al. 2005; Jacquier et al. 2006). This dual agonist–antagonist combinatorial coding is in good agreement with behavioral and psychophysical observations of mixture perception (Laing and Francis 1989; Cometto-Muniz et al. 1999; Spehr et al. 2004; Atanasova et al. 2005; Jacquier et al. 2006).

Deciphering olfactory quality coding thus requires the deorphanization of ORs. Due to the difficulty to functionally express ORs in heterologous cells and the large number of compounds to be tested, identification of odorant repertoires of ORs is laborious. Up to now, ligands were identified for 7 human ORs (Wetzel et al. 1999; Spehr et al. 2003; Matarazzo et al. 2005; Sanz et al. 2005; Jacquier et al. 2006; Saito et al. 2006; Keller et al. 2007). Rodent ORs are more investigated than human ORs and the number of deorphanized rodent ORs is larger (Krautwurst et al. 1998; Touhara et al. 1999; Araneda et al. 2000; Touhara et al. 2000; Gaillard et al. 2002; Oka et al. 2004; Saito et al. 2004; Shirokova et al. 2005; Abaffy et al. 2006; Saito et al. 2006).

Beside the limited number of data concerning human OR odorant repertoires, the sense of smell in humans has long been a subject of studies exploring the relationships between molecular structure and odor qualities (Guillot 1948; Wright 1954; Amoore 1967; Chastrette 1997; Kraft et al. 2000). However, no structure–odor quality theory or model has been proposed to account for the wide range of odors encountered (Laing et al. 2003; Sell 2006). Recent studies of structure–odor quality have attempted to be systematic and have used both homologous and analogous series of odorants. These studies have provided an insight as to how the olfactory system may code odors both at the level of OSN and the various structures within the olfactory bulb and cortex (Imamura et al. 1992; Araneda et al. 2004; Oka et al. 2006; Zou and Buck 2006). A recent work also correlated OSN's responses in Drosophila with structural characteristics of odorant molecules using a quantitative structure–activity relationship (QSAR) approach involving molecular descriptors (Schmuker et al. 2007).

Nevertheless, the link between odor quality of odorants and their detection by ORs is still lacking. Laing et al. (2003) suggested that odorants that share a common quality could reflect activation of a common receptor type. However, working with homologous oxygenated aliphatic molecules, they did not find a common quality to each of the odorants that had the same molecular feature and concluded that identification of the odorants occurs via a combinatorial mechanism involving several types of receptors. However, although the precise odor perception of an odorant should result from a specific combination of activated ORs, it is conceivable that odorant ligands of an OR sharing common structural features could also share a common odor quality. This idea is supported by a recent publication (Doszczak et al. 2007) which describes the possibility to predict the odor of a novel compound on the basis of an OR model. To our point of view, one way to test this hypothesis may be to use OR activation data by odorants instead of odorant homologous series. In this work, we then took advantage of our previous findings concerning the odorant repertoire of the human OR, OR1G1. This receptor was revealed to be broadly tuned, with a preference for 9-carbon aliphatic molecules, and to be inhibited by some odorants structurally related to the agonists (Sanz et al. 2005).

We first performed a 3D-QSAR study of OR1G1 ligands in order to determine the odotope of OR1G1 agonists and antagonists and to predict new ligands for this receptor.

The term ''odotope'' is formed by ''odo,'' which refers to odor, and ''tope'', which refers to topology, describing how spatial features are connected to each other. In this way, odotope meaning joins ''pharmacophore'' meaning: the pharmacophore approach assumes that all the active molecules bind in a common manner to the same target site. The term pharmacophore, first used by Paul Ehrlich in the early 20th century, refers to the ''molecular framework that carries (phoros) the essential features responsible for a drug's (pharmacon) biological activity'' (Ehrlich 1909). The official International Union of Pure and Applied Chemistry definition from 1998 precises: ''a pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response'' and ''the pharmacophore can be considered as the largest common denominator shared by a set of active molecules'' (Wermuth et al. 1998). More recent definitions enhance the crucial role of the spatial arrangement of structural features for a particular biological activity (Beusen and Marshall 2000; Gund 2000). The odotope so appears as the spatial arrangement of chemical features that is recognized by at least one OR, and is linked to an odorant activity.

QSAR studies are less commonly developed in the field of aroma research than in pharmaceutical or toxicological context. Most of the works which were aimed at characterizing odor relationships consisted in classical descriptor 2D-QSAR studies (Anker et al. 1990; Egolf and Jurs 1993; Shvet and Dimoglo 1998; Hadaruga et al. 1999; de Mello Castanho Amboni et al. 2000; Guth et al. 2000, 2001; Wailzer et al. 2001), but several 3D-QSAR investigations have been performed (Frater et al. 1998; Bajgrowicz and Frater 2000; Kraft et al. 2000; Bajgrowicz and Frank 2001; Bajgrowicz et al. 2003). For the current 3D-QSAR study of OR1G1 ligands, we used Catalyst software and a sorting-out procedure of ligands which was successfully used in a previous work and appeared to be advantageously transposable to the present study (Tromelin and Guichard 2003). Predictions provided by the 3D-QSAR approach were experimentally validated. In a second part of this work, we explored the relationships between the molecular features of OR1G1 ligands determined by the 3D-QSAR approach, their activity on the receptor, and their odor quality as perceived by humans.

Materials and methods

Materials and reagents

Odorants were purchased from Sigma-Aldrich, Fluka (Saint Quentin Fallavier, France) or Acros Organics (Noisy-le-Grand, France) at highest purity available. Individual odorants and mixtures were made up fresh by dilutions of stock solutions to final working solution in 100% MeOH (Spectroscopic grade, Sigma). Isoproterenol was purchased from Sigma.

Mammalian expression vectors

As previously described (Sanz et al. 2005), in order to express $G_{\alpha 16}$-protein subunit, we used the pcDNA3.1/Hygro(+) mammalian expression vector (Invitrogen, Cergy-Pontoise, France) carrying the $G_{\alpha 16}$ cDNA (pcDNA3.1/HygroG16). Regarding OR1G1, its gene was introduced into the pCMV-Tag3 mammalian expression vector (Stratagene, Saint-Quentin-en-Yvelines, France), in frame with the c-myc epitope and the first 108 nucleotides of the coding region

of bovine rhodopsin. The resulting vector was named pCMV-RhoTagOR1G1.

Cell culture and transfection of HEK293 cells

HEK293 cells (human embryo kidney cells) were cultured in minimum essential medium without phenol red (GIBCO, Invitrogen corporation, Cergy-Pontoise, France) supplemented with 10% heat-inactivated fetal bovine serum (Hyclone, Perbio, Brebières, France), 2 mM L-glutamine (GIBCO) and Eagle's nonessential amino acids (Eurobio, Les Ulis, France) at 37 °C in a humidified incubator with 5% CO_2. Cells were stably and consecutively transfected with pcDNA3.1/HygroG16 and pCMV-RhoTagOR1G1 plasmids using Lipofectamine 2000 (Invitrogen) according to the manufacturer instructions. HEK293 derivative cells that stably express $G_{\alpha 16}$ or coexpress $G_{\alpha}16$ and OR1G1 were respectively selected in the presence of 300 μg/ml hygromycin B (Invitrogen) or 300 μg/ml hygromycin B and 1 mg/ml neomycin (GIBCO). After selection, cells were frozen in several cryovials in order to use the same cell batches over the study. All cells used were less than 10 passages. OR expression at the cell surface was verified by confocal immunofluorescence microscopy as previously published (Sanz et al. 2005).

Calcium imaging and volatile-odorant functional assay

HEK293 derivative cells were seeded onto a poly-L-lysine–coated 96-well tissue-culture plate (black microtiter plate, Greiner Bio-one, Poitiers, France) at a density of 0.7×10^5 cells per well. Twenty-four hours postseeding, cells were loaded with 2.5 μM of the Ca^{2+}-sensitive fluorescent dye Fluo-4 acetoxymethyl ester (Molecular Probes, Leiden, The Netherlands), as previously described (Sanz et al. 2005). Calcium imaging was carried out at 28 °C using an inverted epifluorescence microscope (CK40 Olympus, Rungis, France) equipped with a digital camera (ORCA-ER, Hamamatsu Photonics, Massy, France). Ca^{2+} reponses were recorded under X10 magnification at 460- to 490-nm excitation and ≥515-nm emission wavelengths. Images were taken every second during 10 min using a bining X2. The SimplePCI software (Hamamatsu, Compix) was used for data acquisition and analysis. The Ca^{2+} signal was expressed as fractional change in fluorescence light intensity: $\Delta F/F = (F - F_0)/F_0$, where F is the fluorescence light intensity at each point and F_0 is the value of emitted fluorescent light before the stimulus application. Cells were counted as responders when $\Delta F/F$ change was at least twice the baseline ΔF/F fluctuation. As previously published (Sanz et al. 2005), OR1G1 response to odorants was expressed as the percentage of responding cells normalized to those responding to the application of 10 μM isoproterenol. Odorants were tested at doses that do not elicit calcium responses from mock-transfected $G_{\alpha 16}$-expressing cells. Because cells were not clonally derived and OR or $G_{\alpha 16}$ expression could vary over time, we also controlled that the number of responding cells to a given odorant (1-nonanol) was constant over the experiments.

To deliver odorants on the cells loaded with Fluo-4, we used volatil odorant functional assay (VOFA) as previously reported (Sanz et al. 2005). When antagonist odorants were tested, agonist and antagonist were mixed into MeOH and coapplied.

Computational methods

Experimental data and molecular building

Activity data used in this work came from a previous study by Sanz et al. (2005). In this study, results from calcium imaging on cells expressing the human OR OR1G1 and stimulated by odorants provided activity values as percent of activated cells. Activity values used for Catalyst hypothesis generation are usually K_d or IC_{50}, so the higher the affinities of ligands, the lower their activity values. In this way, it was essential to transpose the value of percent of activated cells in order to obtain a reliable hierarchical order for Catalyst data. We observed that molecules with different activation levels were similarly classified considering dose-response curves or data at the single dose of 10 μM (see Figure 8 of supplementary material). We chose to use the percent of nonactivated cells reported to the percent of activated cells at 10 μM, and "activity" values were defined as follows: Activity = $\frac{100 - \%\text{cells}}{\%\text{cells}}$. The values are reported in supplementary material.

The 95 compounds tested by Sanz et al. (2005) were built with Catalyst (Catalyst version 4.9.1 software; Accelrys Inc., San Diego, CA, August 2004) running on a Silicon Graphics workstation (SGI-O_2). Catalyst takes into account molecular flexibility by considering each compound as a collection of conformers. The conformers of each compound were generated using the "poling" function (Smellie et al. 1995) of Catalyst/COMPARE module (best conformer generation) to provide the best conformational coverage for a maximum number of conformers generated defaulted to 250 in a 0–20 kcal/mol range from the global minimum.

Hypothesis generation and molecular alignment

HypoGen and HypoRefine modules (Li et al. 2000) were used to perform automated hypothesis generation. In Catalyst, a hypothesis is a model which describes a ligand as a set of chemical functions, defined within catalyst in a dictionary based on atomic characteristics (Greene et al. 1994). These hypotheses should be able to predict the activities of compounds sharing the same receptor-binding mechanism. HypoGen automatically generates the simplest hypotheses that best correlate estimated and experimental affinities. It analyzes the set of "active" molecules first. The most active compound set (usually 5 to a maximum of 8 compounds) is determined using the "uncertainty," noted Unc, so that: $A_{\max} \times Unc - A/Unc > 0$, where $A_{\max}$ is the activity of the most active compound and A the activity of a compound of

the most active set. Starting with the most active molecules, HypoGen analyzes the set of active molecules first, but the hypothesis generation involves the entire set performing a function mapping on each conformer using the function mapping selected. HypoGen selects the best hypothesis by applying a cost analysis (Sutter et al. 2000). The statistical relevance of the various hypotheses is therefore assessed on the basis of their cost relative to the null hypothesis and the fixed hypothesis. The goal of hypothesis generation is to generate a set of hypotheses (defaulted to 10 for a hypothesis generation run) with total costs as close as possible to the fixed cost (Kurogi and Güner 2001). HypoRefine is an extension of HypoGen algorithm that uses exclusion volumes for identifying areas of steric influence on activity. Four chemical functions predefined in the Catalyst Feature Dictionary were used: hydrogen-bond acceptor (HBA), HBA lipid, hydrophobic, and hydrophobic aliphatic. The Unc value is usually defaulted to 3 for hypothesis generation. In the present work, other values were sometimes preferred according to the range of activities of each subset. The defaulted parameters values were used, except for the variable weight parameter which was set to 0.5 (Sutter et al. 2000).

Ligand sorting-out procedure

We use an original procedure to identify several groups of ligands. A previous similar approach, first presented for beta-lactoglobulin ligands (Tromelin and Guichard 2003), was improved in the present context. This procedure involved 3 steps. The first step consisted in dividing the whole initial set into subsets according to the alignment observations: the first group maps all the features of hypothesis model and is called group "A_i"; the second group maps only hydrophobic features and is named group "B_i." Starting from both groups A_i and B_i, the iterative selection procedure occurs as follows: 1) hypothesis generation run, 2) selection of compounds that map all the features of four hypotheses not belonging to the same cluster, and 3) step 1 and 2 are repeated until obtention of a well-ordonned alignment, with satisfactory statistical relevance. The subsets are named "kernel-A" and "kernel-B," respectively. The second step consists in growing starting from the 2 kernel subsets.

Using estimate activities of the initial set of compounds obtained from hypothesis models generated by kernel-A, the well-estimated compounds (the retained criterion was fixed as error <2) were added to constitute subset "A_{n1}," where $n1$ is the number of compounds in the set. With A_{n1}, a hypothesis run was performed, remaining compounds were estimated with generated hypothesis models, and this procedure was repeated until no compound satisfies to fixed criterion error <2. The same protocol is applied to kernel-B. The so finally obtained groups "A_{ni}" and "B_{nj}" possess generally some common compounds. The remaining compounds belonging neither to A_{ni} nor B_{nj} constitute a subset called "Z_i." The third step consists in applying the previous procedure to group "Z_i": identification of "kernel-Z" and growth to obtain a group "Z_{nk}." Compounds of group B_{nj} belong in major part either to group A_{ni} or group Z_{nk} and thus were not considered in the subsequent part of the study.

Validation of provided hypotheses was performed by randomization using CatScramble program available in Catalyst environment and by leave-many-out procedure.

External validation

In order to experimentally validate the modeling, we kept the best significant hypotheses provided by the subsets group A_{ni} and group Z_{nk} (Greene et al. 1994; Kurogi and Güner 2001) and performed a screening by Catalyst database searching through a transposition of the commercial FlavorBase 2004 (Leffingwell & Associates, http://www.leffingwell.com) in Catalyst environment. We then choose a few compounds predicted as strong (activity value lower than 5), medium (activity value between 5 and 19), or weak (activity value higher than 19) agonists and determined their experimental activities as previously described (Sanz et al. 2005).

Odor descriptions

Olfactory descriptions were obtained from the commercial FlavorBase 2004. Descriptors were not ordered and have all the same weight in the analysis.

Statistical analysis

All statistical analyses reported in the odor quality coding part were conducted with SAS software (SAS Institute Inc., Cary, NC), release 8.2.

Correspondence analysis was performed with the CORRESP procedure. Analyses of variance (ANOVA) were performed with the general linear model procedure and corrected P values were calculated using the Bonferroni correction for multiple testing. The stepwise regression with the regression procedure. In the stepwise regression, variables are added one by one to the model, and the F statistic for a variable to be added must be significant at 0.25 level. After a variable is added, however, the stepwise method looks at all the variables already included in the model and deletes any variable that does not produce an F statistic significant at the 0.1 level. Only after this check was made and the necessary deletions accomplished can another variable be added to the model. The stepwise process ends when none of the variables outside the model has an F statistic significant at the 0.25 level and every variable in the model is significant at the 0.1 level or when the variable to be added to the model is the one just deleted from it.

Frequency distribution analyses were performed with the FREQ procedure including 2-sided Fisher's exact test which does not depend on any large sample distribution assumptions and so is appropriate for small sample size tables.

Results

3D-QSAR modeling and experimental validation

Generation of hypothesis models

The hypothesis generation performed on the entire ligand set (list of compounds is available as online supporting information) (95 molecules) revealed 2 compound's alignments as shown in Figure 1. The first one, called group A_i (in dark gray), maps the HBA, and the second one, called group B_i (in light gray), maps only the hydrophobic feature. The whole set was then divided into these 2 groups, and sorting-out procedure was conducted. Starting from group A_i (compounds 1, 3, 4, 7, 10, 15, 17, 19–21, 29, 30, 34, 36, 38, 41–43, 45–48, 50, 56, 58, 59, 61, 63, 69, 73, 74, 77, 80, 81, 86, and 95), the iterative selection procedure based on alignment analysis led to the subset kernel-A that included compounds 1 (1-nonanol), 3 (nonanal), 16 (octanal), 28 (1-heptanol), 37 (decanal), 40 (2-nonanol), 42 (heptanal), 70 (heptanoic acid), and 86 (2-Octanone). The best significant hypothesis is constituted by 2 hydrophobic and 1 HBA features (total cost = 38.0, correlation = 0.98, fixed cost = 28.3, null cost = 197). Starting from kernel-A, growth procedure led to group A35 including 35 compounds (1, 3–5, 10, 17, 19–21, 29, 30, 34, 36, 38, 41–43, 45, 48, 50–52, 56, 58, 59, 61, 64, 74, 75, 78, 81, 84–86, and 89). The best significant hypothesis is constituted by 1 HBA lipid, 1 hydrophobic, and 1 hydrophobic aliphatic features (total cost = 167, correlation = 0.93, fixed cost = 67.8, null cost = 727). All generated hypotheses were validated by randomization and significance value was 99%. Adding methyl salicylate (94) as "inactive compounds," to build group A36, allowed to perform an hypothesis generation run using HypoRefine algorithm. It provided a best significant hypothesis containing an excluded volume in addition to features obtained through HypoGen generation (total cost = 174, correlation = 0.94, fixed cost = 69.4, null cost = 891). Validation was performed by randomization and showed 99% significance for all generated hypotheses. Starting from group B_i (2, 5, 6, 8, 9, 11–14, 22–26, 28, 31, 32, 35, 37, 39, 40, 44, 51, 53–55, 60, 62, 64–68, 70, 71, 72, 75, 76, 78, 79, 82, 83, 85, and 88–94), the iterative selection procedure based on alignment analysis led to the kernel-B that contains only 5 compounds: 27 (ethyl butyrate), 51 (3-octanol), 64 (3-octanone), 78 (butyl butyrate), and 85 (methyl hexanoate). Two hydrophobic and 1 HBA features constituted the best significant hypothesis (total cost = 25.5, correlation = 0.98, fixed cost = 24.1, null cost = 28.9). Growth of kernel-B produced B17, including 17 compounds (25, 27, 37, 45-47, 51, 54, 56, 58, 59, 61, 63, 64, 69, 81, and 94). Note that compounds 45, 51, 56, 58, 59, 61, 64, and 81 belong to both group A35 and B17. Two hydrophobic and 1 HBA features constitute the best significant hypothesis (total cost = 106, correlation = 0.96, fixed cost = 27.1, null cost = 941). Significance was 98% for hypothesis 1–6, 95% for hypothesis 7, 94% for hypothesis 8, 92% for hypothesis 9, and 89% for hypothesis 10. Starting from the 54 compounds not included in A36 and B17 (group Z_i, compounds 2, 6–9, 11–16, 18, 22, 23, 25–28, 32, 33, 35, 37, 39, 40, 44, 46, 47, 49, 53–57, 60, 62, 63, 65–69, 71, 73, 76, 77, 79, 80, 82, 83, 87, 88, 91, 93, and 95) and following the sorting-out alignment procedure, the subset kernel-Z was built with 5 compounds: 2 (ethyl isobutyrate), 11 (methyl thiobutyrate), 40 (vanillin), 55 (dl-camphor), and 7 (cyclohexanone). The best significant hypothesis was based on 2 hydrophobic features and 1 HBA (total cost = 17.4, correlation = 0.999, fixed cost = 17.3, null cost = 87.1). The growth step led to group Z27 (compounds 2, 7, 11, 14–16, 25–27, 31, 37, 40, 44, 54, 55, 66, 67, 70, 71, 76, 79, 82, 83, 87, 88, 90, and 91). Compounds 25, 27, 37, and 54 belong also to the group B17. The best significant hypothesis included 2 hydrophobic features and 1 HBA (total cost = 115, correlation = 0.96, fixed cost = 56.4, null cost = 735). All generated hypotheses were validated by randomization and significance was 99%.

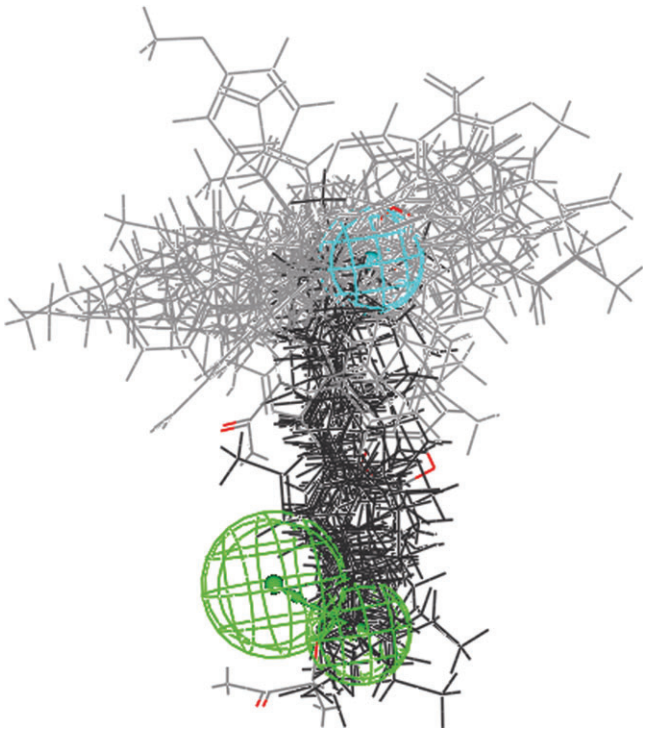

Figure 1 Alignment of the whole training ligand set for the best significant hypothesis. In light blue: hydrophobic features of ligands, corresponding to hydrophobic site on the receptor. In green: HBA features of ligands, constituted by a small sphere corresponding to the center of HBA and a large sphere that is the projection sphere corresponding to a hydrogen-bond donor on the receptor site. In dark gray: group A_i which maps the HBA. In light gray: group B_i which maps only the hydrophobic feature.

Groups A36 and Z27 were kept as reliable ligand groups and were subsequently used as hypothesis models. Corresponding alignments are presented in Figure 2A,B and regression graphs displaying "measured activity" versus "estimated activity" in Figure 2C,D. Orientations of

hydrogen-bond projections and distances between features are displayed in Figure 3. We performed a special leave-many-out validation by removing ''kernel subset'' of each group called in the present case ''leave-kernel-out'' validation. Hypothesis generation run performed on group A26 (compounds 4, 5, 10, 19–21, 30, 34, 36, 42, 45, 48, 50–52, 56, 58, 59, 61, 64, 75, 78, 81, 84, 85, 89, and 94) provided a best significant hypothesis constituted by 1 HBA lipid,

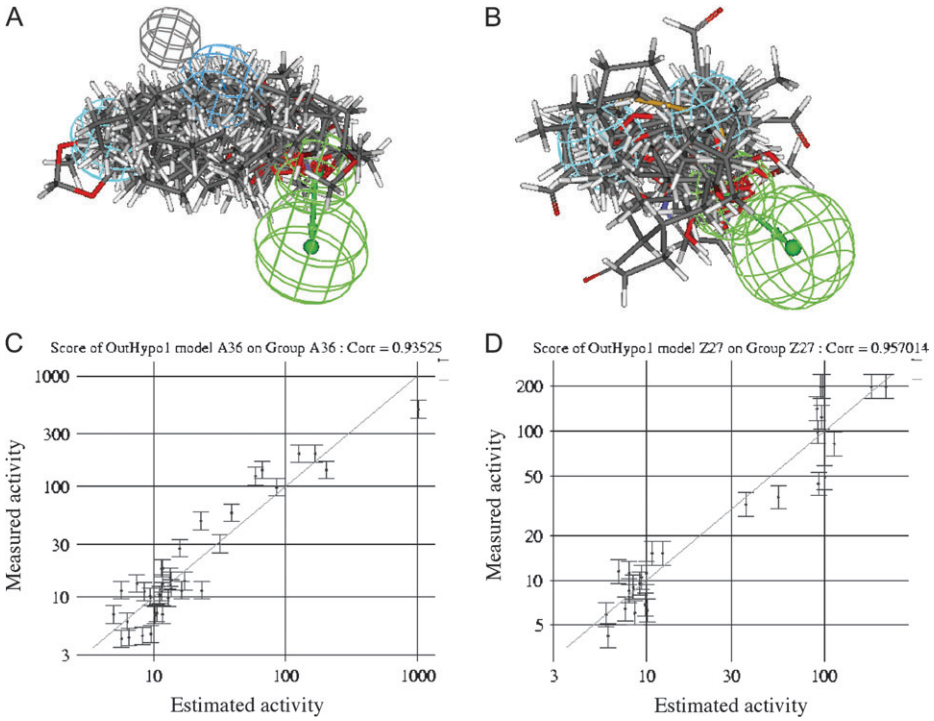

Figure 2 Alignment of groups A36 and Z27 on their respective best significant hypotheses (**A** and **B**, respectively) and corresponding regression graphs displaying ''measured activity'' versus ''estimated activity'' (**C** and **D**, respectively). In dark blue: hydrophobic aliphatic features; in light blue: hydrophobic features. These features correspond to hydrophobic sites on the receptor. In green: HBA features, constituted by a small sphere corresponding to the center of HBA and a large sphere that is the projection sphere corresponding to a hydrogen-bond donor on the receptor site. In gray: excluded volume, indicating a steric hindrance on the receptor site. Estimated activities were obtained by scoring the hypothesis models on their respective training sets.

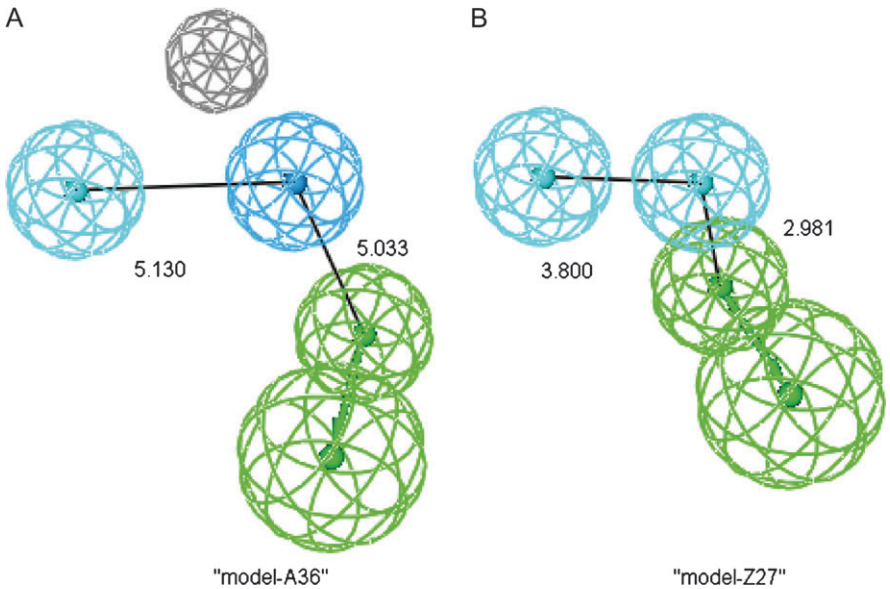

Figure 3 Comparison of ''model-A36'' **(A)** and ''model-Z27'' **(B)**. In dark blue: hydrophobic aliphatic features; in light blue: hydrophobic features; in green: HBA features; in gray: excluded volume (see Figure 2). Distances between centers of features are expressed in Å.

1 hydrophobic, and 1 aliphatic hydrophobic features (total cost = 122, correlation = 0.93, fixed cost = 54.6, null cost = 518). In the same way, hypothesis generation run performed on group Z22 (compounds 7, 14–16, 25–27, 31, 37, 44, 54, 66, 67, 70, 71, 79, 82, 83, 87, 88, 90, and 91) led to a best significant hypothesis constituted by 1 HBA, 1 HBA lipid, and 2 hydrophobic features (total cost = 94.6, correlation = 0.96, fixed cost = total cost = 47.6, null cost = 606). The estimation of activities of "kernel-A" and "kernel-Z," respectively, by both corresponding hypothesis models, provided satisfactory results: correlations between estimated and experimental activities were equal to 0.91 and 0.95, respectively.

Merging groups A and Z

We built a new training set by merging group A36 and group Z27 (compounds 1–5, 7, 10, 11, 14–17, 19–21, 25–27, 29–31, 34, 36–38, 40–45, 48, 50–52, 54–56, 58, 59, 61, 64, 66, 67, 70, 71, 74–76, 78, 79, 81–91, and 94). One HBA lipid, 1 hydrophobic, and 1 aliphatic hydrophobic features make up the best hypothesis characterized by a low significance (cost = 3239, correlation = 0.66, fixed cost = 69.4, null cost = 5701).

Across hypothesis activity estimation

We evaluated activities of group A36 compounds by performing a score of the best significant hypothesis obtained from group Z27 (model-Z27) and conversely a score of the best significant hypothesis obtained from group A36 (model-A36) on group Z27. Whereas, model-A36 did not estimate activities of Z27 compounds with an absolute error value lower than 3, twenty-five A36 compounds were estimated by model-Z27 with an absolute error value lower than 3 (1, 3–5, 10, 17, 19, 20, 21, 29, 30, 34, 36, 38, 41–43, 45, 48, 50–52, 56, 58, and 59).

We tested incorporation of these compounds into group Z27. The best hypothesis provided by hypothesis generation run on the training set so constituted by 62 compounds is formed by 2 hydrophobic features and 1 HBA (cost = 374, correlation = 0.94, fixed cost = 74, null cost = 2543). Addition of twenty-five A36 compounds does not dramatically change models' geometry (data not shown).

In what follows, the reference models and groups A36 and Z27 are called "model-A," "model-Z," "group-A," and "group-Z."

OR1G1 antagonist prediction and validation

In a previous study (Sanz et al. 2005), 3 OR1G1 odorants were demonstrated to inhibit OR1G1 response to several of its agonists. These OR1G1 antagonists were 1-hexanol (31), hexanal (44), and cyclohexanone (76). Our current results show that the 3 compounds belong to group Z and their respective alignments and that of vanillin are close to each others (Figure 4). This suggests that vanillin, which was previously identified as a weak OR1G1 agonist (Sanz et al. 2005), could also be an OR1G1 antagonist. This hypothesis

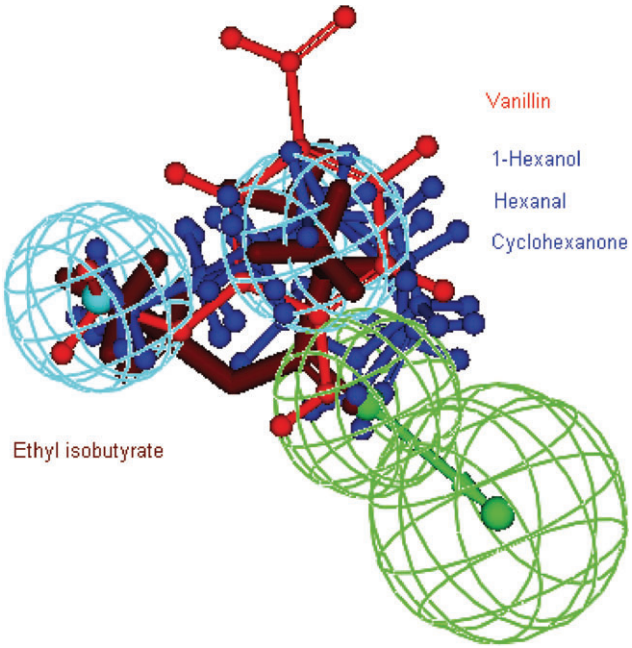

Figure 4 Alignment of the 3 previously identified antagonists, vanillin, and the strong agonist ethyl isobutyrate on "model-Z27." In light blue: hydrophobic features; in green: HBA features (see Figure 1). Vanillin in red; hexanal, hexanol, and cyclohexanone in dark blue; ethyl isobutyrate in brown.

was experimentally tested by heterologous functional expression of OR1G1 in HEK293 cells and calcium imaging, as previously published (Sanz et al. 2005). OR1G1 response to odorants was expressed as percentage of responding cells. Indeed, we have previously demonstrated that, especially in our assays, increased receptor activity is reflected in increased magnitude of the calcium responses of individual cells as well as in increased number of responding cells (Sanz et al. 2005). Results corresponding to cell responses to binary mixtures of vanillin with 3 OR1G1 strong agonists previously identified (Sanz et al. 2005), 1-nonanol, nonanal, and gamma-decalactone, are presented on Figure 5. On one hand, we can again observe that vanillin is a weak agonist of OR1G1 when tested alone. On the other hand, vanillin actually inhibits OR1G1 responses to the 3 tested strong agonists. Vanillin thus appears to be a partial agonist. In order to ensure that the vanillin antagonist effect was OR1G1 specific, vanillin was verified to have no effect on 1) isoproterenol responses of HEK293 cells expressing $G_{\alpha 16}$ (data not shown) and 2) methyl octanoate responses of HEK 293 cells coexpressing $G_{\alpha 16}$ and the receptor OR52D1 which is activated by this odorant (Sanz et al. 2005) (Figure 5). In order to ensure that inhibition by vanillin was not the result of an increased molarity of the mixture compared with 1-nonanol alone, we also tested the effect of another molecule (whiskey lactone) mixed with 10 μM 1-nonanol at doses for which vanillin inhibited OR1G1 responses. OR1G1 response to 10 μM

1-nonanol was not inhibited by whiskey lactone (Figure 5). Both the dose-dependent antagonist effect of vanillin illustrated in Figure 5 and the fact that vanillin is also a weak agonist for OR1G1 suggests a competitive antagonist effect of vanillin. Figure 5 also shows that OR1G1 response to vanillin alone (at a dose exceeding 10 μM) can be greater than OR1G1 response to vanillin mixed with an agonist. This phenomenon could be the result of a competition between vanillin and the agonist which would be unfavorable to the optimal activity of each compound. The possible competitive effect of vanillin is also supported by the fact that OR1G1 antagonists identified until now, including vanillin, belong to the group Z and not to a third group generating a specific hypothesis model. Figure 4 shows that ethyl isobutyrate, which is not an antagonist but the best OR1G1 agonist of group Z, is in the same space as the antagonists, but the position of the carbon chain between the hydrophobic features spheres differs from those of the antagonists.

External validation

In order to validate experimentally the molecular modeling, the best significant hypothesis models A and Z were used to select 10 compounds, predicted as strong (activity value lower than 5), medium (activity value between 5 and 19), or weak (activity higher than 19) agonists or inactive on OR1G1. These newly predicted molecules were tested by OR1G1 functional expression. Experimental results are presented in Figure 6A,B. Estimated activities and experimental activities calculated from results presented in Figure 6 are

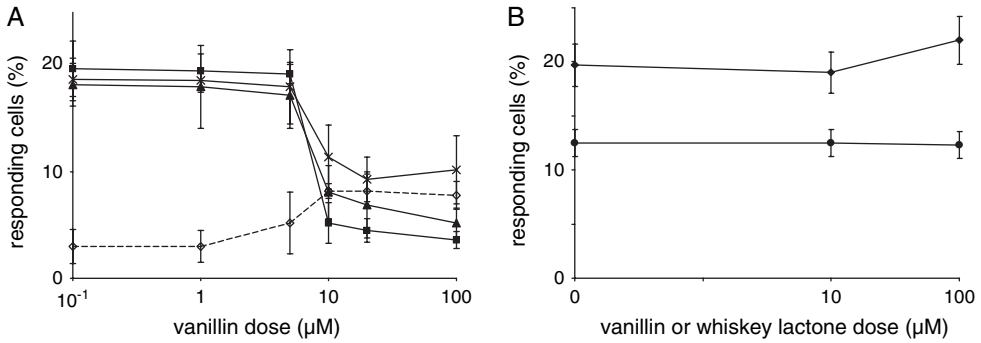

Figure 5 Vanillin antagonism on OR1G1 odorant responses. **(A)** Inhibition of OR1G1 responses to various agonists by vanillin was evaluated using VOFA and HEK293 cells coexpressing OR1G1 and $G_{\alpha 16}$. Vanillin was tested alone (◇ and dotted line) or mixed with 10 μM of nonanal (×), γ-decalactone (▲), or 1-nonanol (■). Vanillin doses varied from 0.1 to 100 μM. Ca^{2+} responses were recorded during 10 min and data are shown as number of responding cells. Bars indicate standard deviation (SD; 3 independent experiments). **(B)** Effect of vanillin, at doses for which vanillin inhibits OR1G1 responses (10–100 μM), on responses to methyl octanoate of HEK293 cells coexpressing $G_{\alpha 16}$ and OR52D1. (●) 10 μM methyl octanoate mixed with vanillin. Effect of whiskey lactone, at doses for which vanillin inhibits OR1G1 responses (10–100 μM), on OR1G1 response to 10 mM 1-nonanol. (♦) 10 μM 1-nonanol mixed with whiskey lactone. Bars indicate SD.

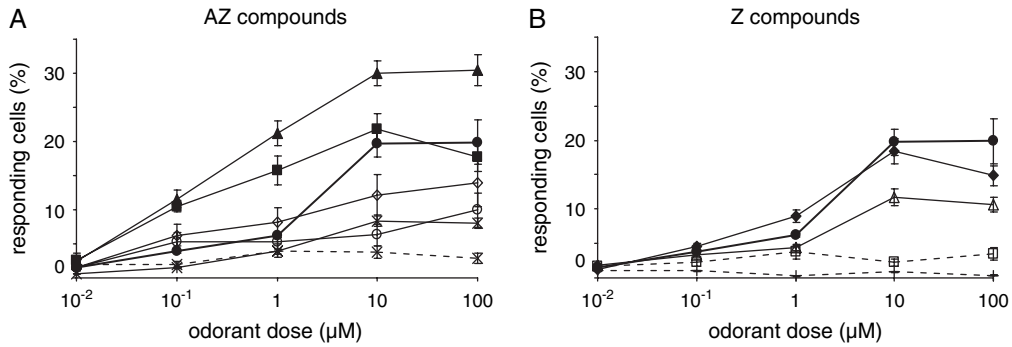

Figure 6 OR1G1 responses to predicted agonists. Different odorants predicted as OR1G1 agonists were tested using VOFA and HEK293 cells coexpressing OR1G1 and $G_{\alpha 16}$. Odorants were applied at 0.01 to 100 μM. Ca^{2+} responses were recorded during 10 min and reported as the number of responding cells normalized as percentage of cells responding to 10 μM isoproterenol. At the end of each experiment, isoproterenol (10 μM) was applied to verify the ability of cells to generate Ca^{2+} responses via the $G_{\alpha 16}$-protein subunit. $G_{\alpha 16}$-expressing HEK293 cells were also stimulated with each of the tested odorants at 100 μM in order to control the specificity of the response obtained in presence of OR1G1 (data not shown). Bars indicate standard deviation (3 independent experiments). **(A)** AZ compounds (● and thick line) 1-nonanol (OR1G1 agonist previously identified and used as a reference), (v) 9-decen-1-ol, (▲) tridecanal, (◇) 2-methylundecanal, (○) 3-methylthio-1-hexanol, (× and dotted line) 5-hydroxy-4-octanone, (*) *trans*-2-*trans*-4-nonadienal. **(B)** (■ and thick line) 1-nonanol as a reference and Z compounds, (□ and dotted line) *cis*-4-hexen-1-ol, (Δ) 4-phenyl-butan-1-ol, (♦) 3-methyl-1-pentanol, and (+ and dotted line) *n*-butanal.

reported in Table 1. Tridecanal, 9-decen-1-ol, and 3-methyl-1-pentanol, predicted as strong agonists, were revealed to be more or equally active on OR1G1 than 1-nonanol, a previously identified potent agonist for OR1G1. 2-methylundecanal, 4-phenyl-butan-1-ol, and *trans*-2-*trans*-4-nonadienal appeared to be medium agonists as predicted. Yet, 3-methylthio-1-hexanol, predicted as medium agonist, was demonstrated as a weak agonist, and 5-hydroxy-4-octanone and *cis*-4-hexen-1-ol, predicted as weak agonists, in fact did not succeed in activating OR1G1. The predicted inactive compound, *n*-butanal, was indeed unable to activate OR1G1. So, whatever the hypothesis model, ''A'' or ''Z,'' we can observe that agonists predicted as strong or medium (tridecanal, 9-decen-1-ol, 3-methyl-1-pentanol, 2-methylundecanal, 4-Phenyl-butan-1-ol, *trans*-2-*trans*-4-nonadienal, and 3-methylthio-1-hexanol) were actually able to activate OR1G1. Nevertheless, some estimated activities were not totally in agreement with experimental data, but improvement of estimation should be obtained by hypothesis generation on new trainings sets formed by incorporation of compounds ''AZ type'' in training sets A or Z and for compounds ''Z type'' in training set Z (data not shown). From our results, we can also observe that 2 molecules, 3-methyl-1-pentanol and 4-phenyl-butan-1-ol, are exclusively predicted by the model ''Z,'' that demonstrates the relevance of the 2 models. The correlation between the predicted and measured activity was tested by regression analysis using estimated activity according to compound type (A or Z) reported in Table 1 (the used values are underlined). Estimated activity values obtained by model-A were used for 2-methylundecanal, *trans*-2-*trans*-4-nonadienal, and 3-methylthio-1-hexanol, whereas estimated activity values obtained by model-Z were used for tridecanal, 9-decen-1-ol, 3-methyl-1-pentanol, 4-phenylbutan-1-ol, 5-hydroxy-4-octanone, *cis*-4-hexen-1-ol, and *n*-butanal. Regression parameters (experimental activities vs. estimated activities) are reported in Table 2. Very close statistical values were obtained using the activity value estimated by model-A for 9-decen-1-ol (r^2 = 0.9477, F test = 144.89).

We also tested predictive power of model obtained by merging of groups A36 and Z27. Estimated activities and experimental activities are reported in Table 1, whereas regression parameters are reported Table 3.

Odor quality coding

On the basis of the above structure–activity relationships of OR1G1 odorant ligands, we attempted to correlate odor quality of these odorants with their molecular structure and biological activity on OR1G1.

Odor quality descriptions of OR1G1 best agonists seem very similar. Indeed, descriptors such as oily, fatty, floral, or rose were repeatedly found in the quality description of the 5 agonists with the highest experimentally measured activity (Table 4). In order to confirm this observation, the 105 odorants tested on OR1G1 were split up into 3 groups (on the basis of a cluster analysis, data not shown) depending on their activity: 8 odorants in the strong group (activity < 5.0), 58 in the medium group (5.0 < activity < 19.0) and 39 in the weak group (19.0 < activity). The correspondence between experimental activities values (in vitro activities) and olfactory description was then explored through a correspondence analysis (McEwan and Schlich 1991/92). To perform this analysis, only 13 descriptors with an occurrence higher than

Table 1 Estimated and experimental activity values[a] of new tested compounds

Compounds	Estimated activity values by hypothesis "model-A"	Estimated activity values by hypothesis "model-Z"	Estimated activity values by hypothesis "merge AZ model"	Experimental activity values	Compound type
Tridecanal	6.5	5.9	11	2.3	AZ
9-Decen-1-ol	5.2	5.1	10	3.6	AZ
3-Methyl-1-pentanol	370	5.3	12	4.4	Z
2-Methylundecanal	6.9	5	11	7.3	AZ
4-Phenyl-butan-1-ol	530	10	59	7.5	Z
trans-2, *trans*-4-nonadienal	9.3	14	11	10.9	AZ
3-Methylthio-1-hexanol	8.2	5.4	8	14.8	AZ
5-Hydroxy-4-octanone	83	8.4	10	25.6	AZ
cis-4-Hexen-1-ol	540	22	68	43.4	Z
n-Butanal	640	92	71	113.3	Z (inactive)

Activities were estimated using the best significant hypothesis models A, Z, and AZ, respectively, obtained by automated generation of hypotheses for group A-36, Z-27, and merged AZ. Underlined values were used for regression analysis in Table 2.
[a]Activity = $\frac{100 - \%\text{cells}}{\%\text{cells}}$.

Table 2 Regression parameters of correlation between experimental and estimated activity values by models A or Z (experimental values vs. estimated values)

Statistical regression analysis					
r^2	0.9478				
Standard error	8.2488				
F	145.21				
F critical	5.32				
Observations	10				
Variance analysis					
	Degree of freedom	Sum of squares			
Regression	1	9890.64	9880.64		
Residuals	8	544.35	68.04		
Total	9	10424.98			
	Coefficient	Standard error	*t* Statistic	Probability	CI (α = 5%) min value; max value
Intercept point	1.8409	3.1597	0.5826	0.5762	−5.4454; 9.1272
Slope	1.24138	0.1030	12.0503	2.08E-06	1.0038; 1.4789

Used values for regression analysis are underlined in columns "estimated activity values by Hypothesis model-A" and "estimated activity values by Hypothesis model-Z" in Table 1.

Table 3 Regression parameters of correlation between experimental and estimated activity values by model merge AZ model (experimental values vs. estimated values)

Statistical regression analysis					
r^2	0.4963				
Standard error	25.5544				
F	7.8818				
F critical	5.3176				
Observations	10				
Variance analysis					
	Degree of freedom	Sum of squares	Mean of squares		
Regression	1	5147.01	5147.01		
Residuals	8	5224.20	653.03		
Total	9	10371.21			
	Coefficient	Standard error	*t* Statistic	Probability	CI (α = 5%) min value; max value
Intercept point	−0.6651	11.7587	−0.0566	0.9563	−27.7807; 26.4506
Slope	0.8849	0.3152	2.8075	0.0229	0.1581; 1.6118

Used values for regression analysis are reported in column "estimated activity values by merge AZ model" in Table 1.

5% were kept (cheese, cognac, fatty, floral, fruity, green, herbaceous, nutty, oily, rose, sweet, waxy, and winey). Figure 7 represents a combined plot derived from correspondence analysis between groups of OR1G1 agonists (based on in vitro activity) and their odor description. On this map, the closeness of a group (weak, medium, or strong) and an odor descriptor is interpreted in terms of correspondence. The further away 2 close points are from the origin, the stronger the implications. It was concluded from this multivariate analysis that there was a correspondence between a high in vitro

agonist activity (strong group) and descriptors such as rose and waxy. This observation suggested a link between in vitro activity values and odor description, subsequently revealing the possible implication of OR1G1 in olfactory coding of these odor notes.

To evaluate the statistical significance of such a link, we performed an ANOVA between the 13 odor descriptors of OR1G1 agonists and their in vitro measured activity (n = 105). The model was not found to be significant ($F_{13,104}$ = 1.2; P = 0.29). Therefore, we performed ANOVAs between each of the 13 odor descriptors of agonists and their in vitro measured activity. ANOVA models were found to be significant for rose ($F_{1,104}$ = 10.6, P = 0.002; corrected P = 0.021) and to a lesser extent for waxy ($F_{1,104}$ = 5.5, P = 0.02; corrected P = 0.27) descriptors. In order to investigate whether a combination of factors would lead to a better model, we conducted a stepwise regression between odor descriptor occurrence and in vitro activity. Only the 2 previously identified descriptors were found to be positively and significantly (α = 0.1) linked to activity: rose (P = 0.006) and waxy (P = 0.09).

Moreover, we observed that from the set of the 105 odorants tested on OR1G1, 75% of the strong agonists (Table 4) are AZ type odorants and 83% of the strong AZ type OR1G1 agonists carry rose, waxy, or fatty odor. Furthermore, from the total tested set, there is no strong OR1G1 agonist which has a rose, waxy, or fatty odor and is not of AZ type. Only 4-phenyl-butan-1-ol, which is a medium agonist of Z type exclusively, has a rose odor quality. Hence, it seems that the best OR1G1 agonists are in majority of AZ type and have a rose, waxy, or fatty odor quality. It is then likely that rose waxy or fatty notes are essentially linked with strong and AZ type OR1G1 ligands. Statistical analyses of the olfactory description of AZ type OR1G1 ligands as compared with non-AZ type OR1G1 ligands (2-sided Fisher's exact test [FET]) confirmed these observations. In all, 48% of the AZ type OR1G1 ligands were described as fatty (P = 0.0002, corrected P = 0.003, FET), 26% as waxy (P = 0.005, corrected P = 0.07, FET), and 19% as rose (P = 0.008, corrected P = 0.10, FET).

Table 4 Experimental activity values[a] and odor descriptions of best OR1G1 agonists

Compounds	Experimental activity values[a]	Odor description	Compound Type
Tridecanal	2.3	Powerful waxy, citrus odor, fatty	AZ
9-Decen-1-ol	3.6	Waxy, fatty, rosy	AZ
1-Nonanol	4.2	Oily floral, rose orange	AZ
Ethyl isobutyrate	4.3	Sweet, ethereal, fruity	Z
Nonanal	4.4	Fatty floral rose, waxy	AZ
3-Methyl-1-pentanol	4.4	Winey, cognac, whisky, fruity, green	Z
γ-Decalactone	4.5	Coconut peach	AZ
2-Ethyl-1-hexanol	4.7	Sweet, oily, weak rose	AZ

[a]Activity = $\frac{100 - \%\text{cells}}{\%\text{cells}}$.

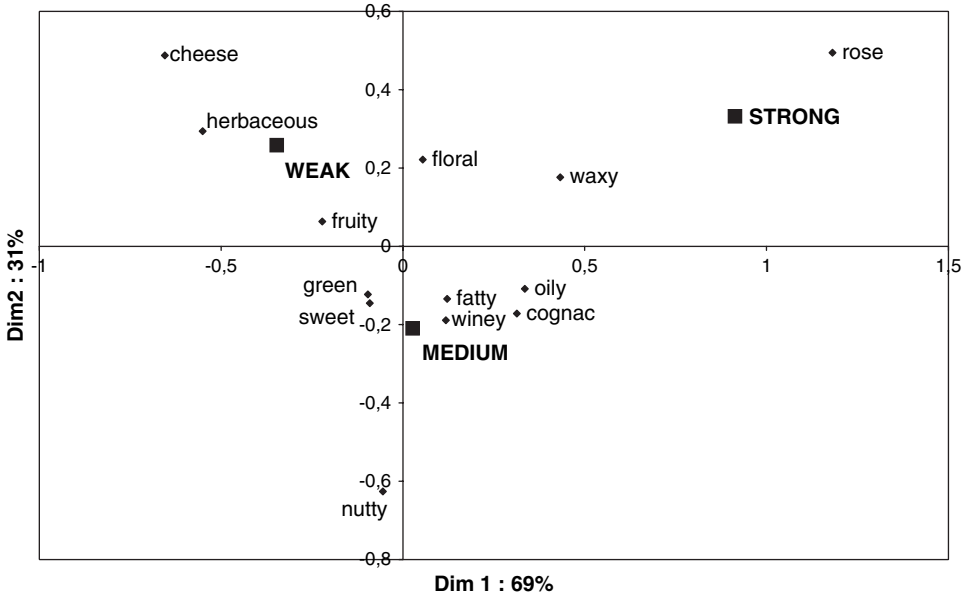

Figure 7 Combined plot derived from correspondence analysis between groups of OR1G1 agonists (based on in vitro activity) and their odor description. The closeness of a group (weak, medium, or strong) and an odor descriptor is interpreted in terms of correspondence. For clarity reasons, all descriptors were not reported on the map. Strong group: activity < 5.0, medium group: 5.0 < activity < 19.0, weak group: 19.0 < activity.

Discussion

The present study investigated the relationships between molecular structure, in vitro agonistic performance and odor description of odorants interacting with the human OR OR1G1. Starting from in vitro biological data, 3D-QSAR approach showed that hypotheses generated on the whole set of 95 compounds failed to reveal a single model for OR1G1 ligands. This suggests that OR1G1 could exhibit several binding modes. Indeed, our sorting-out selection of compounds led to 2 subsets, which permitted to obtain highly significant hypothesis models, so called ''model-A'' (for the best significant hypothesis provided by group A36) and ''model-Z,'' respectively(for the best significant hypothesis provided by group Z27). As a common point, 1 HBA (or HBA lipid) and 2 hydrophobic features make up these 2 models. However, orientation of hydrogen-bond projections is different and distances between features are smaller for model-Z, as shown in Figure 3. When increasing the spacing parameter to a value equal or higher than 4 Å (data not shown), generation of hypotheses by group Z27 is not possible. This small distance was also observed for hypotheses provided by group ''leave-kernel-out Z22,'' confirming the crucial aspect of this geometry. Conversely, theses changes of distance and tolerance parameters have no, or only little effect, on generated hypothesis models provided by group A36, which is characterized by an excluded volume that puts forward the presence of a steric hindrance in the receptor site.

Indeed, hypothesis generation run performed on the training set constituted by association of compounds A36 and Z27 produced poor statistical significant model, emphasizing the specificity of each model. However, model-Z was able to correctly estimate activities of several compounds belonging to group A36, whereas there was no correct estimation performed by model-A on compounds Z27. This property should be related to distance between features presented before: mapping most of the molecules on model-Z was facilitated by short distance between spheres features of model-Z, but conversely mapping small molecules on hypothesis model-A was disadvantaged, even impossible.

The predictive power of these models was then attested by the experimental external validation of 10 compounds previously not tested. Moreover, the 3D-QSAR approach allowed us to propose some associations of ligands and restricted the choices of their conformers, which could greatly facilitate the further docking of OR1G1 ligands through molecular modeling of the receptor-binding site and suggests clues for the understanding of odorant perception.

Having demonstrated that odorant ligands of OR1G1 share common structural features, we addressed the question whether these ligands could also share a common odor quality. When taking into account the 105 odorants tested with OR1G1, we observed a correspondence between odorant activity on OR1G1 and olfactory description. ANOVA and stepwise regression also suggest that OR1G1 ligands would be associated with rose odor and to a lesser extent with waxy odor. Nevertheless, this needs to be validated with external data. Taking into account the 3D-QSAR categorization of OR1G1 ligands, we further showed by statistical analyses of the olfactory description that AZ type OR1G1 ligands, which correspond to 75% of the strong agonists, are linked with fatty, waxy, and rose notes. Hence, it is worth noticing that the best OR1G1 ligands appear in majority to have a rose, waxy, or fatty odor and to be of AZ type. These results thus suggest that an odotope could be associated to an odor quality. Obviously, it could be that the distinct odor quality of OR1G1 agonists groups found in the present study was due to the interaction of these agonists with at least another OR. Furthermore, although our data suggest that the way an OR binds structurally related odorants might partly determine their odor quality, the perception of each odorant remains unique and probably results from the activation of a specific combination of ORs.

Along the olfactory pathway, it is known that monomolecular odorants are represented in the main olfactory bulb by distinct spatial patterns of activated glomeruli, corresponding to specific groups of activated ORs (Ressler et al. 1994; Vassar et al. 1994; Mombaerts et al. 1996). Moreover, individual glomeruli responds to odorants sharing a specific combination of molecular features and odorants structurally related activate neighboring glomeruli (Imamura et al. 1992; Xu et al. 2003; Mori et al. 2006; Oka et al. 2006). Nevertheless, structurally related odorants may have different perceived odor (Laing et al. 2003). In mixture studies, it was reported that, at the level of the olfactory bulb, glomeruli activated by a mixture of odorants correspond to the sum of the glomeruli activated by its components (Laing et al. 2003; Lin et al. 2006). Yet, it was shown that cortical neurons activated by a mixture of 2 odorants are not stimulated by individual odorants (Zou and Buck 2006). This result is in line with psychophysical studies demonstrating that perceived odor of an odorant mixture does not correspond to the superposition of odor quality of its components (Jinks and Laing 2001). Hence, although structure of odorants seems encoded at the level of the olfactory bulb, it is lacking a code of odor quality. A recent work by Gottfried et al. (2006) showed that, in the piriform cortex, codes of odorant structure and odor quality are dissociated, but the mechanism by which the odor quality information in piriform cortex is integrated with the odorant structure code remains unknown. Aside, our work put ORs at the crossroads, integrating both structure and quality informations about odorants.

Recent studies (Schmiedeberg et al. 2007; Stary et al. 2007) demonstrated that ORs with high homology (orthologues such as human OR1A1 and mouse Olfr43 or paralogues such as human OR1A1 and human OR1A2) bind common ligands with similar efficacy, whereas ORs more distantly related (such as Olfr43 and Olfr49) binds common odorants but with different efficacies. The study by Schmiedeberg

et al. also demonstrated that evolutionary conserved amino acid positions define the ligand-binding site. Then, these studies suggest that an odorant would be recognized via a similar odotope by closely related ORs, whereas distantly related ORs would bind a common ligand via different odotopes. In the present work, we demonstrate that an OR can recognize 2 odotopes, suggesting that the binding pocket of an OR can accommodate several odotopes. When taken together, our results also suggest that odorants sharing a same odotope recognized by OR1G1 would evoke similar odor quality. So, an odotope could be associated with an odor quality. But, as an odorant can be recognized by various ORs via different odotopes, it can be suggested that odotopes recognized with the best efficacies would define the prominent perceived odor quality.

In summary, the findings reported here provide a new insight in the understanding of the relationships between odorants, ORs and odor quality. It has especially been shown for the first time by a 3D-QSAR approach that ligands of an OR, OR1G1, have to be divided in 2 groups in order to find satisfactory models, suggesting 2 modes of interaction of odorants with this receptor. This result is in agreement with another study (Sell 2006) reporting that it would be difficult to design a model for a typical ligand for OR1G1. In another part of this work, we also reported the likely involvement of OR1G1 in the perception of waxy, fatty, and rose odor in humans. These results support the idea that, among the specific group of ORs activated by an odorant and defining its particular odor perception, some ORs that strongly bind this odorant might determine its major odor quality. Similar studies for other ORs need to be made in order to validate this view and to increase the understanding and predictability of odor perception.

Supplementary material

Supplementary material can be found at http://www.chemse.oxfordjournals.org/

Funding

Institut National de la Recherche Agronomique (France); the Agence Nationale de la Recherche (AROMALIM program); by the Région Ile de France (SESAME grant number A01947).

Acknowledgements

The authors would like to thank C. Chabanet for statistical advices and Dr. P. Etiévant and the AROMAGRI research group for early discussions on this project.

References

Abaffy T, Matsunami H, Luetje CW. 2006. Functional analysis of a mammalian odorant receptor subfamily. J Neurochem. 97:1506–1518.

Amoore JE. 1967. Specific anosmia: a clue to the olfactory code. Nature. 214:1095–1098.

Anker LS, Jurs PC, Edwards PA. 1990. Quantitative structure-retention relationship studies of odor-active aliphatic compounds with oxygen-containing functional groups. Anal Chem. 62:2676–2684.

Araneda RC, Kini AD, Firestein S. 2000. The molecular receptive range of an odorant receptor. Nat Neurosci. 3:1248–1255.

Araneda RC, Peterlin Z, Zhang X, Chesler A, Firestein S. 2004. A pharmacological profile of the aldehyde receptor repertoire in rat olfactory epithelium. J Physiol. 555:743–756.

Atanasova B, Thomas-Danguin T, Chabanet C, Langlois D, Nicklaus S, Etiévant PX. 2005. Perceptual interactions in odour mixtures: odour quality in binary mixtures of woody and fruity wine odorants. Chem Senses. 30:209–217.

Bajgrowicz JA, Berg-Schultz K, Brunner G. 2003. Substituted hepta-1,6-dien-3-ones with green/fruity odours green/galbanum olfactophore model. Bioorg Med Chem. 11:2931–2946.

Bajgrowicz JA, Frank I. 2001. Camphor-derived amber/woody odorants: 1,7,7-trimethyl-2′-iso-propylspiro[bicyclo[2.2.1]heptane-2,4′-(1,3-dioxanes)]. Tetrahedron Asymmetry. 12:2049–2057.

Bajgrowicz JA, Frater G. 2000. Chiral recognition of sandalwood odorants. Enantiomer. 5:225–234.

Beusen DD, Marshall GR. 2000. Pharmacophore definition using the active analog approach. In: Güner OF, editor. Pharmacophore perception, development and use in drug design. La Jolla (CA): International University Line. p. 21–46.

Chastrette M. 1997. Trends in structure-odor relationships. SAR QSAR Environ Res. 6:215–254.

Cometto-Muniz JE, Cain WS, Abraham MH, Gola JMR. 1999. Chemosensory detectability of 1-butanol and 2-heptanone singly and in binary mixtures. Physiol Behav. 67:269–276.

de Mello Castanho Amboni RD, da Silva Junkes B, Yunes RA, Fonseca Heinzen VE. 2000. Quantitative structure-odor relationships of aliphatic esters using topological indices. J Agric Food Chem. 48:3517–3521.

Doszczak L, Kraft P, Weber HP, Bertermann R, Triller A, Hatt H, Tacke R. 2007. Prediction of perception: probing the hOR17-4 olfactory receptor model with silicon analogues of bourgeonal and lilial. Angew Chem Int Ed Engl. 46:3367–3371.

Duchamp-Viret P, Chaput MA, Duchamp A. 1999. Odor response properties of rat olfactory receptor neurons. Science. 284:2171–2174.

Duchamp-Viret P, Duchamp A, Chaput MA. 2003. Single olfactory sensory neurons simultaneously integrate the components of an odour mixture. Eur J Neurosci. 18:2690–2696.

Egolf LM, Jurs PC. 1993. Quantitative structure-retention and structure-odor intensity relationships for a diverse group of odor-active compounds. Anal Chem. 65:3119–3126.

Ehrlich P. 1909. Present status of chemotherapy. Ber Dtsch Chem Ges. 42:17–47.

Frater G, Bajgrowicz JA, Kraft P. 1998. Fragrance Chemistry. Tetrahedron. 54:7633–7703.

Gaillard I, Rouquier S, Pin JP, Mollard P, Richard S, Barnabe C, Demaille J, Giorgi D. 2002. A single olfactory receptor specifically binds a set of odorant molecules. Eur J Neurosci. 15:409–418.

Glusman G, Yanai I, Rubin I, Lancet D. 2001. The complete human olfactory subgenome. Genome Res. 11:685–702.

Gottfried JA, Winston JS, Dolan RJ. 2006. Dissociable codes of odor quality and odorant structure in human piriform cortex. Neuron. 49:467–479.

Greene J, Kahn SD, Savoj H, Sprague PW. 1994. Chemical function queries for 3D database search. J Chem Inf Comput Sci. 34:1297–1308.

Guillot M. 1948. Anosmies partielles et odeurs fondamentales. C R Acad Sci. 226:1307–1309.

Gund P. 2000. Evolution of the pharmacophore concept in pharmaceutical research. In: Güner OF, editor. Pharmacophore perception, development and use in drug design. La Jolla (CA): International University Line. p. 3–12.

Guth H, Buhr K, Fritzler R. 2000. Relationship between odour activity and molecular structure of lactones. In: Schieberle P, Engel KH, editors. Frontiers in flavour science. Garching (Germany): Deutsche Forschungsanstalt für Lebensmittelchemie. p. 235–242.

Guth H, Buhr K, Fritzler R. 2001. Descriptors for structure-property correlation studies of odorants. In: Takeoka GR, Güntert M, Engel KH, editors. Aroma active compounds in foods. Chemistry and sensory properties. Washington, DC: American Chemical Society. p. 93–108.

Hadaruga DI, Muresan S, Bologa C, Chiriac A, Simon Z, Cofar L, Naray-Szabo G. 1999. QSAR for cycloaliphatic alcohols with qualitatively defined sandalwood odour characteristics. QSAR. 18:253–261.

Imamura K, Mataga N, Mori K. 1992. Coding of odor molecules by mitral tufted cells in rabbit olfactory-bulb .1. Aliphatic-compounds. J Neurophysiol. 68:1986–2002.

Jacquier V, Pick H, Vogel H. 2006. Characterization of an extended receptive ligand repertoire of the human olfactory receptor OR17-40 comprising structurally related compounds. J Neurochem. 97:537–544.

Jinks A, Laing DG. 2001. The analysis of odor mixtures by humans: evidence for a configurational process. Physiol Behav. 72:51–63.

Keller A, Zhuang HY, Chi QY, Vosshall LB, Matsunami H. 2007. Genetic variation in a human odorant receptor alters odour perception. Nature. 449:468–472.

Kraft P, Bajgrowicz JA, Denis C, Fráter G. 2000. Odds and trends: recent developments in the chemistry of odorants. Angew Chem Int Ed Engl. 39:2981–3010.

Krautwurst D, Yau KW, Reed RR. 1998. Identification of ligands for olfactory receptors by functional expression of a receptor library. Cell. 95: 917–926.

Kurogi Y, Güner OF. 2001. Pharmacophore modeling and three-dimensional database searching for drug design using catalyst. Curr Med Chem. 8:1035–1055.

Laing DG, Francis GW. 1989. The capacity of humans to identify odors in mixtures. Physiol Behav. 46:809–814.

Laing DG, Legha PK, Jinks AL, Hutchinson I. 2003. Relationship between molecular structure, concentration and odor qualities of oxygenated aliphatic molecules. Chem Senses. 28:57–69.

Li H, Sutter J, Hoffman R. 2000. HypoGen: an automated system for generating 3D predictive pharmacophore models. In: Güner OF, editor. Pharmacophore perception, development and use in drug design. La Jolla (CA): IUL Biotechnlogy Series. p. 171–189.

Lin DY, Shea SD, Katz LC. 2006. Representation of natural stimuli in the rodent main olfactory bulb. Neuron. 50:937–949.

Malnic B, Godfrey PA, Buck LB. 2004. The human olfactory receptor gene family. Proc Natl Acad Sci USA. 101:2584–2589.

Malnic B, Hirono J, Sato T, Buck LB. 1999. Combinatorial receptor codes for odors. Cell. 96:713–723.

Matarazzo V, Clot-Faybesse O, Marcet B, Guiraudie-Capraz G, Atanasova B, Devauchelle G, Cerutti M, Etievant P, Ronin C. 2005. Functional characterization of two human olfactory receptors expressed in the baculovirus Sf9 insect cell system. Chem Senses. 30:195–207.

McEwan JA, Schlich P. 1991/92. Correspondence analysis in sensory evaluation. Food Qual Prefer. 3:23–36.

Mombaerts P, Wang F, Dulac C, Chao SK, Nemes A, Mendelsohn M, Edmondson J, Axel R. 1996. Visualizing an olfactory sensory map. Cell. 87:675–686.

Mori K, Takahashi YK, Igarashi KM, Yamaguchi M. 2006. Maps of odorant molecular features in the mammalian olfactory bulb. Physiol Rev. 86:409–433.

Niimura Y, Nei M. 2005. Evolutionary dynamics of olfactory receptor genes in fishes and tetrapods. Proc Natl Acad Sci USA. 102:6039–6044.

Niimura Y, Nei M. 2007. Extensive gains and losses of olfactory receptor genes in mammalian evolution. PLoS ONE. 2:e708.

Oka Y, Katada S, Omura M, Suwa M, Yoshihara Y, Touhara K. 2006. Odorant receptor map in the mouse olfactory bulb: In vivo sensitivity and specificity of receptor-defined glomeruli. Neuron. 52:857–869.

Oka Y, Omura M, Kataoka H, Touhara K. 2004. Olfactory receptor antagonism between odorants. Embo J. 23:120–126.

Ressler KJ, Sullivan SL, Buck LB. 1994. Information coding in the olfactory system - evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell. 79:1245–1255.

Saito H, Chi Q, Zhuang H, Matsunami H. 2006. The functional properties of mammalian odorant receptors. Chem Senses. 31:A65.

Saito H, Kubota M, Roberts RW, Chi QY, Matsunami H. 2004. RTP family members induce functional expression of mammalian odorant receptors. Cell. 119:679–691.

Sanz G, Schlegel C, Pernollet JC, Briand L. 2005. Comparison of odorant specificity of two human olfactory receptors from different phylogenetic classes and evidence for antagonism. Chem Senses. 30:69–80.

Schmiedeberg K, Shirokova E, Weber HP, Schilling B, Meyerhof W, Krautwurst D. 2007. Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2. J Struct Biol. 159:400–412.

Schmuker M, de Bruyne M, Hahnel M, Schneider G. 2007. Predicting olfactory receptor neuron responses from odorant structure. Chem Cent J. 1:11.

Sell CS. 2006. On the unpredictability of odor. Angew Chem Int Ed Engl. 45:6254–6261.

Shirokova E, Schmiedeberg K, Bedner P, Niessen H, Willecke K, Raguse JD, Meyerhof W, Krautwurst D. 2005. Identification of specific ligands for orphan olfactory receptors—G protein-dependent agonism and antagonisms of odorants. J Biol Chem. 280:11807–11815.

Shvet NM, Dimoglo AS. 1998. Structure-odour relationships: results of an applied electron-topological approach. Nahrung. 42:364–370.

Smellie A, Teig SL, Tobwin P. 1995. Poling: promoting conformational variation. J Comput Chem. 16:171–187.

Spehr M, Gisselmann G, Poplawski A, Riffell JA, Wetzel CH, Zimmer RK, Hatt H. 2003. Identification of a testicular odorant receptor mediating human sperm chemotaxis. Science. 299:2054–2058.

Spehr M, Schwane K, Heilmann S, Gisselmann G, Hummel T, Hatt H. 2004. Dual capacity of a human olfactory receptor. Curr Biol. 14:R832–R833.

Stary A, Suwattanasophon C, Woschann P, Buchbauer G. 2007. Differences in (-)citronellal binding to various odorant receptors. Biochem Biophys Res Commun. 361:941–945.

Sutter J, Güner OF, Hoffman R. 2000. Effects of variable weights and tolerances on predictive model generation. In: Güner OF, editor. Pharmacophore perception, development and use in drug design. La Jolla (CA): IUL Biotechnlogy Series. p. 2988–3003.

Touhara K, Sengoku S, Inaki K, Tsuboi A, Hirono J, Sato T, Sakano H, Haga T. 1999. Functional identification and reconstitution of an odorant receptor in single olfactory neurons. Proc Natl Acad Sci USA. 96: 4040–4045.

Touhara K, Sengoku S, Inaki K, Tsuboi A, Hirono J, Sato T, Sakano H, Haga T. 2000. Functional identification and reconstitution of an odorant receptor in single olfactory neurons. . Proc Natl Acad Sci USA. 97:3782[vol. 96, p. 4040, 1999].

Tromelin A, Guichard E. 2003. Use of catalyst in a 3D-QSAR study of the interactions between flavor compounds and β-lactoglobulin. J Agric Food Chem. 51:1977–1983.

Vassar R, Chao SK, Sitcheran R, Nunez JM, Vosshall LB, Axel R. 1994. Topographic organization of sensory projections to the olfactory bulb. Cell. 79:981–991.

Wailzer B, Klocker J, Buchbauer G, Ecker G, Wolschann P. 2001. Prediction of the aroma quality and the threshold values of some pyrazines using artificial neural networks. J Med Chem. 44:2805–2813.

Wermuth G, Ganellin CR, Lindberg P, Mitscher LA. 1998. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). Pure Appl Chem. 70:1129–1143.

Wetzel CH, Oles M, Wellerdieck C, Kuczkowiak M, Gisselmann G, Hatt H. 1999. Specificity and sensitivity of a human olfactory receptor functionally expressed in human embryonic kidney 293 cells and Xenopus laevis oocytes. J Neurosci. 19:7426–7433.

Wright RH. 1954. Odour and molecular vibration. 1. Quantum and thermodynamic considerations. J Appl Chem. 4:611–615.

Xu FQ, Liu N, Kida L, Rothman DL, Hyder F, Shepherd GM. 2003. Odor maps of aldehydes and esters revealed by functional MRI in the glomerular layer of the mouse olfactory bulb. Proc Natl Acad Sci USA. 100:11029–11034.

Zou ZH, Buck LB. 2006. Combinatorial effects of odorant mixes in olfactory cortex. Science. 311:1477–1481.

Zozulya S, Echeverri F, Nguyen T. 2001. The human olfactory receptor repertoire. Genome Biol. 2:0018.0011–0018.0012.

Accepted June 4, 2008